



Amdt C

SEQUENCE LISTING

<110> Fisher, Paul B.
Kang, Dong-Chul
Gopalkrishnan, Rahul V.

<120> MELANOMA DIFFERENTIATION ASSOCIATED
GENE-5 AND PROMOTER AND USES THEREOF

<130> A34614 (070050.1690)

<140> 09/515,363

<141> 2000-02-29

<160> 25

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 3365

<212> DNA

<213> homo sapiens

<400> 1

gcgcgcggc ctgagagccc tggacaac ctcgtcattt tcaggcacag agcgtagac 60
cctgcttc taagtggca gggacagcg gcacgcacat ttcacctgtc ccgcacacaa 120
cagcaccatc tgcttggag aaccctctcc cttctctgag aaagaaagat gtcgaatggg 180
tattccacag acgagaattt ccgtatctc atctcggtct tcagggccag ggtaaaaatg 240
tacatccagg tggagcctgt gctggactac ctgaccccttc tgcctgcaga ggtgaaggag 300
cagattcaga ggacagtgcg cacccggg aacatgcagg cagttgaact gctgctgagc 360
accttggaga agggagtctg gcacccctgg tggactcggg aattcgtgaa ggcctccgg 420
agaaccggca gcccctggc cggccgtac atgaaccctg agtcacgaa cttggccctct 480
ccatcgttt agaacgtctca tggatatat ctccaactgc tgaacccctt tcagcccact 540
ctgggtggaca agcttcttagt tagagacgctc ttggataagt gcatggagga ggaactgttg 600
acaattgaag acagaaaccg gattgctgct gcagaaaaca atggaaatgaa atcaggtgta 660
agagagctac taaaaaggat tggcagaaa gaaaactggt tctctgcatt tctgaatgtt 720
cttcgtcaaa caggaaacaa tggacttgc caagagttt caggctctgaa ttgctcagaa 780
agcaatgcag agattgagaa ttatcacaa gttgatggc ctcaagtggaa agagcaactt 840
cttcaacca cagttcagcc aaatctggag aaggaggctt gggcatggaa gaataactca 900
tcagaatcat ctttgcaga ttcttctgtt gttcagaat cagacacaag tttggcagaa 960
ggaagtgtca gctgctttaga tggaaagtctt ggacataaca gcaacatggg cagtgattca 1020
ggcaccatgg gaagtgattc agatgaagag aatgtggcag caagagcatc cccggagcca 1080
gaactccagc tcaggccta ccaaattggaa gttgcccagc cagccttggaa agggaaagaat 1140
atcatcatct gcctccctac agggagtggaa aaaaccagag tggctttt cattgccaag 1200
gatcacttag acaagaagaa aaaagcatct gagcctggaa aagttatagt tcttgtcaat 1260
aagttactgc tagttgaaca gctctccgc aaggagttcc aaccatttt gaagaaatgg 1320
tatcggttta ttggattaag tggatacc caactgaaaa tatcattttc agaagttgtc 1380
aagtccctgt atattattat cagtagct caaatccctt aaaaactccct cttaaacttg 1440
aaaaatggag aagatgctgg tggcaattt tcagactttt ccctcattat cattgatgaa 1500
tgtcatcaca ccaacaaaga agcagtgtat aataacatca tgaggcatta tttgatgcag 1560
aagttgaaaa acaatagact caagaaagaa aacaaaccag tgattccct tcctcagata 1620
ctgggactaa cagttcacc tggatggaa gggccacga agcaagccaa agctgaagaa 1680
cacattttaa aactatgtgc caatcttgat gcatttacta taaaactgt taaagaaaac 1740
cttgatcaac tggaaaacca aatacaggag ccatgaaaga agtttgcacat tgcaatgc 1800

accagagaag atccatttaa agagaaactt ctagaaataa tgacaaggat tcaaacttat 1860
tgtcaatga gtccaatgtc agatttgga actcaaccct atgaacaatg ggccattcaa 1920
atgaaaaaaa aagctgcaaa aaaaggaaat cgcaaagaac gtgttgcgaga agaacattt 1980
aggaagtaca atgaggccct acaaattaat gacacaattc gaatgataga tgcgtatact 2040
catctgaaa ctittataa tgaagagaaa gataagaatg ttgcagtcat agaagatgat 2100
agtatgagg gtggatgatga tgatgttgcgatg aagatgagga tgattaaag 2160
aaaccttga aactggatga aacagataga tttctcatga cttttatttt tgaaaacaat 2220
aaaatgttga aaaggctggc tggaaaccca gaatatgaaa atgaaaagct gaccaaatta 2280
agaaatacca taatggagca atatactagg actgaggaat cagcacgagg aataatctt 2340
acaaaaacac gacagagtgc atatgcgctt tcccagtggta ttactgaaa tgaaaaattt 2400
gctgaagtag gactcaaaagc ccaccatctg attggagctg gacacagcag tgagttcaaa 2460
cccatgacac agaatgaaca aaaagaagtc attagtaat ttgcactgg aaaaatcaat 2520
ctgcttatcg ctaccacagt ggccagaagaa ggtctggata ttaaagaatg taacattgtt 2580
atccgttatg gtctcgac caatgaaata gccatggtc aggccctgg tcgagccaga 2640
gctgatgaga gcacctacgt cctgggtgcgatg cccatgggtt caggagttt cgaacatgag 2700
acagttatg atttccgaga gaagatgatg tataaagcta tacattgtt tcaaaatatg 2760
aaaccagagg agtatgctca taagattttg gaattacaga tgcaaagtat aatggaaaag 2820
aaaatgaaaa ccaagagaaa tattgccaag cattacaaga ataaccatc actaataact 2880
ttcccttgca aaaactgcag tggcttagcc tggctgggg aagatatcca tgtaatttag 2940
aaaatgcac acgtcaatat gaccccgagaa ttcaaggaac ttacattgtt aagagaaaac 3000
aaagcactgc aaaagaagtgc tgccgactat caaataatg gtgaaatcat ctgcaaattgt 3060
ggccaggctt ggggaacaat gatggtgcac aaaggcttag atttgccttg tctcaaaata 3120
aggaattttg tagtgggtttt caaaaataat tcaacaaaga aacaatacaa aaagtggta 3180
gaattaccta tcacatttcc caatcttgac tattcagaat gctgtttatt tagtgatgag 3240
gattagcaact tgattgaaga ttctttaaa atactatcg ttaaacatatt aatatgatta 3300
tgattaatgt attcattatg ctacagaact gacataagaa tcaataaaat gattgtttt 3360
ctctg 3365

<210> 2
<211> 1025
<212> PRT
<213> homo sapiens

<400> 2
Met Ser Asn Gly Tyr Ser Thr Asp Glu Asn Phe Arg Tyr Leu Ile Ser
1 5 10 15
Cys Phe Arg Ala Arg Val Lys Met Tyr Ile Gln Val Glu Pro Val Leu
20 25 30
Asp Tyr Leu Thr Phe Leu Pro Ala Glu Val Lys Glu Gln Ile Gln Arg
35 40 45
Thr Val Ala Thr Ser Gly Asn Met Gln Ala Val Glu Leu Leu Leu Ser
50 55 60
Thr Leu Glu Lys Gly Val Trp His Leu Gly Trp Thr Arg Glu Phe Val
65 70 75 80
Glu Ala Leu Arg Arg Thr Gly Ser Pro Leu Ala Ala Arg Tyr Met Asn
85 90 95
Pro Glu Leu Thr Asp Leu Pro Ser Pro Ser Phe Glu Asn Ala His Asp
100 105 110
Glu Tyr Leu Gln Leu Leu Asn Leu Leu Gln Pro Thr Leu Val Asp Lys
115 120 125
Leu Leu Val Arg Asp Val Leu Asp Lys Cys Met Glu Glu Leu Leu
130 135 140
Thr Ile Glu Asp Arg Asn Arg Ile Ala Ala Ala Glu Asn Asn Gly Asn
145 150 155 160
Glu Ser Gly Val Arg Glu Leu Leu Lys Arg Ile Val Gln Lys Glu Asn
165 170 175
Trp Phe Ser Ala Phe Leu Asn Val Leu Arg Gln Thr Gly Asn Asn Glu

180	185	190
Leu Val Gln Glu Leu Thr Gly	Ser Asp Cys Ser Glu Ser Asn Ala Glu	
195	200	205
Ile Glu Asn Leu Ser Gln Val Asp Gly Pro Gln Val Glu Glu Gln Leu		
210	215	220
Leu Ser Thr Thr Val Gln Pro Asn Leu Glu Lys Glu Val Trp Gly Met		
225	230	235
Glu Asn Asn Ser Ser Glu Ser Ser Phe Ala Asp Ser Ser Val Val Ser		
245	250	255
Glu Ser Asp Thr Ser Leu Ala Glu Gly Ser Val Ser Cys Leu Asp Glu		
260	265	270
Ser Leu Gly His Asn Ser Asn Met Gly Ser Asp Ser Gly Thr Met Gly		
275	280	285
Ser Asp Ser Asp Glu Glu Asn Val Ala Ala Arg Ala Ser Pro Glu Pro		
290	295	300
Glu Leu Gln Leu Arg Pro Tyr Gln Met Glu Val Ala Gln Pro Ala Leu		
305	310	315
Glu Gly Lys Asn Ile Ile Ile Cys Leu Pro Thr Gly Ser Gly Lys Thr		
325	330	335
Arg Val Ala Val Tyr Ile Ala Lys Asp His Leu Asp Lys Lys Lys Lys		
340	345	350
Ala Ser Glu Pro Gly Lys Val Ile Val Leu Val Asn Lys Val Leu Leu		
355	360	365
Val Glu Gln Leu Phe Arg Lys Glu Phe Gln Pro Phe Leu Lys Lys Trp		
370	375	380
Tyr Arg Val Ile Gly Leu Ser Gly Asp Thr Gln Leu Lys Ile Ser Phe		
385	390	395
Pro Glu Val Val Lys Ser Cys Asp Ile Ile Ile Ser Thr Ala Gln Ile		
405	410	415
Leu Glu Asn Ser Leu Leu Asn Leu Glu Asn Gly Glu Asp Ala Gly Val		
420	425	430
Gln Leu Ser Asp Phe Ser Leu Ile Ile Ile Asp Glu Cys His His Thr		
435	440	445
Asn Lys Glu Ala Val Tyr Asn Asn Ile Met Arg His Tyr Leu Met Gln		
450	455	460
Lys Leu Lys Asn Asn Arg Leu Lys Lys Glu Asn Lys Pro Val Ile Pro		
465	470	475
Leu Pro Gln Ile Leu Gly Leu Thr Ala Ser Pro Gly Val Gly Gly Ala		
485	490	495
Thr Lys Gln Ala Lys Ala Glu Glu His Ile Leu Lys Leu Cys Ala Asn		
500	505	510
Leu Asp Ala Phe Thr Ile Lys Thr Val Lys Glu Asn Leu Asp Gln Leu		
515	520	525
Lys Asn Gln Ile Gln Glu Pro Cys Lys Lys Phe Ala Ile Ala Asp Ala		
530	535	540
Thr Arg Glu Asp Pro Phe Lys Glu Lys Leu Leu Glu Ile Met Thr Arg		
545	550	555
Ile Gln Thr Tyr Cys Gln Met Ser Pro Met Ser Asp Phe Gly Thr Gln		
565	570	575
Pro Tyr Glu Gln Trp Ala Ile Gln Met Glu Lys Lys Ala Ala Lys Lys		
580	585	590
Gly Asn Arg Lys Glu Arg Val Cys Ala Glu His Leu Arg Lys Tyr Asn		
595	600	605
Glu Ala Leu Gln Ile Asn Asp Thr Ile Arg Met Ile Asp Ala Tyr Thr		
610	615	620
His Leu Glu Thr Phe Tyr Asn Glu Glu Lys Asp Lys Lys Phe Ala Val		
625	630	635
		640

Ile Glu Asp Asp Ser Asp Asp Glu Gly Gly Asp Asp Glu Tyr Cys Asp Gly
645 650 655
Asp Glu Asp Glu Asp Asp Leu Lys Lys Pro Leu Lys Leu Asp Glu Thr
660 665 670
Asp Arg Phe Leu Met Thr Leu Phe Phe Glu Asn Asn Lys Met Leu Lys
675 680 685
Arg Leu Ala Glu Asn Pro Glu Tyr Glu Asn Glu Lys Leu Thr Lys Leu
690 695 700
Arg Asn Thr Ile Met Glu Gln Tyr Thr Arg Thr Glu Glu Ser Ala Arg
705 710 715 720
Gly Ile Ile Phe Thr Lys Thr Arg Gln Ser Ala Tyr Ala Leu Ser Gln
725 730 735
Trp Ile Thr Glu Asn Glu Lys Phe Ala Glu Val Gly Val Lys Ala His
740 745 750
His Leu Ile Gly Ala Gly His Ser Ser Glu Phe Lys Pro Met Thr Gln
755 760 765
Asn Glu Gln Lys Glu Val Ile Ser Lys Phe Arg Thr Gly Lys Ile Asn
770 775 780
Leu Leu Ile Ala Thr Thr Val Ala Glu Glu Gly Leu Asp Ile Lys Glu
785 790 795 800
Cys Asn Ile Val Ile Arg Tyr Gly Leu Val Thr Asn Glu Ile Ala Met
805 810 815
Val Gln Ala Arg Gly Arg Ala Arg Ala Asp Glu Ser Thr Tyr Val Leu
820 825 830
Val Ala His Ser Gly Ser Gly Val Ile Glu His Glu Thr Val Asn Asp
835 840 845
Phe Arg Glu Lys Met Met Tyr Lys Ala Ile His Cys Val Gln Asn Met
850 855 860
Lys Pro Glu Glu Tyr Ala His Lys Ile Leu Glu Leu Gln Met Gln Ser
865 870 875 880
Ile Met Glu Lys Lys Met Lys Thr Lys Arg Asn Ile Ala Lys His Tyr
885 890 895
Lys Asn Asn Pro Ser Leu Ile Thr Phe Leu Cys Lys Asn Cys Ser Val
900 905 910
Leu Ala Cys Ser Gly Glu Asp Ile His Val Ile Glu Lys Met His His
915 920 925
Val Asn Met Thr Pro Glu Phe Lys Glu Leu Tyr Ile Val Arg Glu Asn
930 935 940
Lys Ala Leu Gln Lys Lys Cys Ala Asp Tyr Gln Ile Asn Gly Glu Ile
945 950 955 960
Ile Cys Lys Cys Gly Gln Ala Trp Gly Thr Met Met Val His Lys Gly
965 970 975
Leu Asp Leu Pro Cys Leu Lys Ile Arg Asn Phe Val Val Val Phe Lys
980 985 990
Asn Asn Ser Thr Lys Lys Gln Tyr Lys Lys Trp Val Glu Leu Pro Ile
995 1000 1005
Thr Phe Pro Asn Leu Asp Tyr Ser Glu Cys Cys Leu Phe Ser Asp Glu
1010 1015 1020
Asp
1025

<210> 3
<211> 1036
<212> DNA
<213> homo sapiens

<220>
<221> misc_feature
<222> 551
<223> n = A,T,C or G

<400> 3
gcacatttg gcctacaaag gaccttattt ttaaggcaga acctgctggg aaaacaaaat 60
atccgccgga ggagcttgt agagcgttgg tcttgggtgc agagagaatt cgcttcctt 120
ttctgtttcc cgccgggttcc ttaaccaaag gcctcctctc ttcacccggc cccgaccaaaa 180
ggtggcgtct ccctgaggaa actccctccc cgccaggcag attacgttta caaagtctt 240
agaagagaat cgaaacagaa accaaagtca ggcaaactct gtaagaactg cctgacagaa 300
agctggactc aaagctctta cccgagtgta cagcaggatc gccccgggtcc gggacccccc 360
gcgcacacccg cagagtccaa agtgcgcgcg ctgccccccg cacctgcctg cccggcccc 420
gcgcgcccgc cccgctgccc cctgccccccg tgcccacctg cccaggtgcg agtgcagccc 480
cgccgcgcgg cctgagagcc ctgtggacaa cctcgctatt gtcaggcaca gagcggtaga 540
ccctgcttct ntaagtgggc agcggacagg ggcacgcaca tttcacctgt cccgcagaca 600
acagcaccat ctgcttggga gaaccctctc ccttcctctga gaaagaaaaga tgcataatgg 660
gtattccaca gacgagaatt tccgctatct catctcgatc ttcagggcca gggtaaaaat 720
gtacatccag gtggagcctg tgctggacta cctgaccttt ctgcctgcag aggtgaagga 780
gcagattcag aggacagtcg ccacctccgg gaacatgcag gcagttgaac tgctgctgag 840
caccttggag aaggagatct ggacacccggg ttggactcgg gaattcgtgg aggccctccg 900
gagaaccggc agccctctgg ccgcggccta catgaaccct gagtcacgg acttgcctc 960
tccatcgttt gagaacgctc atgatgaata tctccaactg ctgaacccctt ttcagccac 1020
tctgggtggac aagctt 1036

<210> 4
<211> 202
<212> DNA
<213> homo sapiens

<220>
<221> misc_feature
<222> 33
<223> n = A,T,C or G

<221> misc_feature
<222> 73
<223> n = A,T,C or G

<221> misc_feature
<222> 107
<223> n = A,T,C or G

<400> 4
cattgaactc ttttaagaa cacaatatac tangcattat ccattttattt gttggcaga 60
ggtaaggaaa atntaccaat aattttcatc agtgtggagc attatantcc tgtggaaaga 120
atgctgaagt acaaatgaga atccaaagta ccagtcgttcc ttctgtcact aatttcaga 180
ataaaattag gcaaatcagt tc 202

<210> 5
<211> 112
<212> PRT
<213> homo sapiens

<400> 5
Met Ala Glu Gly Asn His Arg Lys Lys Pro Leu Lys Val Leu Glu Ser
1 5 10 15

Leu Gly Lys Asp Phe Leu Thr Gly Val Leu Asp Asn Leu Val Glu Gln
20 25 30
Asn Val Leu Asn Trp Lys Glu Glu Lys Lys Tyr Tyr Asp Ala
35 40 45
Lys Thr Glu Asp Lys Val Arg Val Met Ala Asp Ser Met Gln Glu Lys
50 55 60
Gln Arg Met Ala Gly Gln Met Leu Leu Gln Thr Phe Phe Asn Ile Asp
65 70 75 80
Gln Ile Ser Pro Asn Lys Lys Ala His Pro Asn Met Glu Ala Gly Pro
85 90 95
Pro Glu Ser Gly Glu Ser Thr Asp Ala Leu Lys Leu Cys Pro His Glu
100 105 110

<210> 6
<211> 118
<212> PRT
<213> homo sapiens

<400> 6
Met Ala Asp Lys Val Leu Lys Glu Lys Arg Lys Leu Phe Ile Arg Ser
1 5 10 15
Met Gly Glu Gly Thr Ile Asn Gly Leu Leu Asp Glu Leu Leu Gln Thr
20 25 30
Arg Val Leu Asn Lys Glu Glu Met Glu Lys Val Lys Arg Glu Asn Ala
35 40 45
Thr Val Met Asp Lys Thr Arg Ala Leu Ile Asp Ser Val Ile Pro Lys
50 55 60
Gly Ala Gln Ala Cys Gln Ile Cys Ile Thr Tyr Ile Cys Glu Glu Asp
65 70 75 80
Ser Tyr Leu Ala Gly Thr Leu Gly Leu Ser Ala Ala Pro Gln Ala Val
85 90 95
Gln Asp Asn Pro Ala Met Pro Thr Ser Ser Gly Ser Glu Gly Asn Val
100 105 110
Lys Leu Cys Ser Leu Glu
115

<210> 7
<211> 108
<212> PRT
<213> homo sapiens

<400> 7
Met Ala Ser Asp Asp Leu Ser Leu Ile Arg Lys Asn Arg Met Ala Leu
1 5 10 15
Phe Gln Gln Leu Thr Cys Val Leu Pro Ile Leu Asp Asn Leu Leu Lys
20 25 30
Ala Asn Val Ile Asn Lys Gln Glu His Asp Ile Ile Lys Gln Lys Thr
35 40 45
Gln Ile Pro Leu Gln Ala Arg Glu Leu Ile Asp Thr Ile Leu Val Lys
50 55 60
Gly Asn Ala Ala Ala Asn Ile Phe Lys Asn Cys Leu Lys Glu Ile Asp
65 70 75 80
Ser Thr Leu Tyr Lys Asn Leu Phe Val Asp Lys Asn Met Lys Tyr Ile
85 90 95
Pro Thr Glu Asp Val Ser Gly Leu Ser Leu Glu Glu

100

105

<210> 8
<211> 107
<212> PRT
<213> homo sapiens

<400> 8
Glu Ser Asn Asp Leu Leu Leu Ile Arg Lys Asn Arg Met Ala Leu Phe
1 5 10 15
Gln His Leu Thr Cys Val Ile Pro Ile Leu Asp Ser Leu Leu Thr Ala
20 25 30
Gly Ile Ile Asn Glu Gln Glu His Asp Val Ile Lys Gln Lys Thr Gln
35 40 45
Thr Ser Leu Gln Ala Arg Glu Leu Ile Asp Thr Ile Leu Val Lys Gly
50 55 60
Asn Ile Ala Ala Thr Val Phe Arg Asn Ser Leu Gln Glu Ala Glu Ala
65 70 75 80
Val Leu Tyr Glu His Leu Phe Val Gln Gln Asp Ile Lys Tyr Ile Pro
85 90 95
Thr Glu Asp Val Ser Asp Leu Pro Val Glu Glu
100 105

<210> 9
<211> 115
<212> PRT
<213> homo sapiens

<400> 9
Met His Pro His His Gln Glu Thr Leu Lys Lys Asn Arg Val Val Leu
1 5 10 15
Ala Lys Gln Leu Leu Leu Ser Glu Leu Leu Glu His Leu Leu Glu Lys
20 25 30
Asp Ile Ile Thr Leu Glu Met Arg Glu Leu Ile Gln Ala Lys Val Gly
35 40 45
Ser Phe Ser Gln Asn Val Glu Leu Leu Asn Leu Leu Pro Lys Arg Gly
50 55 60
Pro Gln Ala Phe Asp Ala Phe Cys Glu Ala Leu Arg Glu Thr Lys Gln
65 70 75 80
Gly His Leu Glu Asp Met Leu Leu Thr Thr Leu Ser Gly Leu Gln His
85 90 95
Val Leu Pro Pro Leu Ser Cys Asp Tyr Asp Leu Ser Leu Pro Phe Pro
100 105 110
Val Cys Glu
115

<210> 10
<211> 50
<212> PRT
<213> homo sapiens

<400> 10
Leu Val Asp Lys Leu Leu Val Arg Asp Val Leu Asp Lys Cys Met Glu
1 5 10 15

Glu Glu Leu Leu Thr Ile Glu Asp Arg Asn Arg Ile Ala Ala Ala Glu
20 25 30
Asn Asn Gly Asn Glu Ser Gly Val Arg Glu Leu Leu Lys Arg Ile Val
35 40 45
Gln Lys
50

<210> 11
<211> 109
<212> PRT
<213> homo sapiens

<400> 11
Met Glu Ala Arg Asp Lys Gln Val Leu Arg Ser Leu Arg Leu Glu Leu
1 5 10 15
Gly Ala Glu Val Leu Val Glu Gly Leu Val Leu Gln Tyr Leu Tyr Gln
20 25 30
Glu Gly Ile Leu Thr Glu Asn His Ile Gln Glu Ile Asn Ala Gln Thr
35 40 45
Thr Gly Leu Arg Lys Thr Met Leu Leu Asp Ile Leu Pro Ser Arg
50 55 60
Gly Pro Lys Ala Phe Asp Thr Phe Leu Asp Ser Leu Gln Glu Phe Pro
65 70 75 80
Trp Val Arg Glu Lys Leu Lys Lys Ala Arg Glu Glu Ala Met Thr Asp
85 90 95
Leu Pro Ala Gly Asp Arg Leu Thr Gly Ile Pro Ser His
100 105

<210> 12
<211> 406
<212> PRT
<213> mus musculus

<400> 12
Met Ser Ala Ser Gln Asp Ser Arg Ser Arg Asp Asn Gly Pro Asp Gly
1 5 10 15
Met Glu Pro Glu Gly Val Ile Glu Ser Asn Trp Asn Glu Ile Val Asp
20 25 30
Ser Phe Asp Asp Met Asn Leu Ser Glu Ser Leu Leu Arg Gly Ile Tyr
35 40 45
Ala Tyr Gly Phe Glu Lys Pro Ser Ala Ile Gln Gln Arg Ala Ile Leu
50 55 60
Pro Cys Ile Lys Gly Tyr Asp Val Ile Ala Gln Ala Gln Ser Gly Thr
65 70 75 80
Gly Lys Thr Ala Thr Phe Ala Ile Ser Ile Leu Gln Gln Ile Glu Leu
85 90 95
Asp Leu Lys Ala Thr Gln Ala Leu Val Leu Ala Pro Thr Arg Glu Leu
100 105 110
Ala Gln Gln Ile Gln Lys Val Val Met Ala Leu Gly Asp Tyr Met Gly
115 120 125
Ala Ser Cys His Ala Cys Ile Gly Gly Thr Asn Val Arg Ala Glu Val
130 135 140
Gln Lys Leu Gln Met Glu Ala Pro His Ile Ile Val Gly Thr Pro Gly
145 150 155 160
Arg Val Phe Asp Met Leu Asn Arg Arg Tyr Leu Ser Pro Lys Tyr Ile

165	170	175	
Lys Met Phe Val Leu Asp Glu Ala Asp Glu Met Leu Ser Arg Gly Phe			
180	185	190	
Lys Asp Gln Ile Tyr Asp Ile Phe Gln Lys Leu Asn Ser Asn Thr Gln			
195	200	205	
Val Val Leu Leu Ser Ala Thr Met Pro Ser Asp Val Leu Glu Val Thr			
210	215	220	
Lys Lys Phe Met Arg Asp Pro Ile Arg Ile Leu Val Lys Lys Glu Glu			
225	230	235	240
Leu Thr Leu Glu Gly Ile Arg Gln Phe Tyr Ile Asn Val Glu Arg Glu			
245	250	255	
Glu Trp Lys Leu Asp Thr Leu Cys Asp Leu Tyr Glu Thr Leu Thr Ile			
260	265	270	
Thr Gln Ala Val Ile Phe Ile Asn Thr Arg Arg Lys Val Asp Trp Leu			
275	280	285	
Thr Glu Lys Met His Ala Arg Asp Phe Thr Val Ser Ala Met His Gly			
290	295	300	
Asp Met Asp Gln Lys Glu Arg Asp Val Ile Met Arg Glu Phe Arg Ser			
305	310	315	320
Gly Ser Ser Arg Val Leu Ile Thr Thr Asp Leu Leu Ala Arg Gly Ile			
325	330	335	
Asp Val Gln Gln Val Ser Leu Val Ile Asn Tyr Asp Leu Pro Thr Asn			
340	345	350	
Arg Glu Asn Tyr Ile His Arg Ile Gly Arg Gly Arg Phe Gly Arg			
355	360	365	
Lys Gly Val Ala Ile Asn Met Val Thr Glu Glu Asp Lys Arg Thr Leu			
370	375	380	
Arg Asp Ile Glu Thr Phe Tyr Asn Thr Ser Ile Glu Glu Met Pro Leu			
385	390	395	400
Asn Val Ala Asp Leu Ile			
405			

<210> 13
 <211> 500
 <212> PRT
 <213> molgula oculata

<400> 13			
Met Ser Gly Tyr Ser Ser Asp Arg Asp Arg Gly Arg Asp Arg Gly Phe			
1	5	10	15
Gly Ala Pro Arg Phe Gly Gly Ser Arg Ala Gly Pro Leu Ser Gly Lys			
20	25	30	
Lys Phe Gly Asn Pro Gly Glu Lys Leu Val Lys Lys Lys Trp Asn Leu			
35	40	45	
Asp Glu Leu Pro Lys Phe Glu Lys Asn Phe Tyr Gln Glu His Pro Asp			
50	55	60	
Leu Ala Arg Arg Thr Ala Gln Glu Val Glu Thr Tyr Arg Arg Ser Lys			
65	70	75	80
Glu Ile Thr Val Arg Gly His Asn Cys Pro Lys Pro Val Leu Asn Phe			
85	90	95	
Tyr Glu Ala Asn Phe Pro Ala Asn Val Met Asp Val Ile Ala Arg Gln			
100	105	110	
Asn Phe Thr Glu Pro Thr Ala Ile Gln Ala Gln Gly Trp Pro Val Ala			
115	120	125	
Leu Ser Gly Leu Asp Met Val Gly Val Ala Gln Thr Gly Ser Gly Lys			
130	135	140	

Thr Leu Ser Tyr Leu Leu Pro Ala Ile Val His Ile Asn His Gln Pro
145 150 155 160
Phe Leu Glu Arg Gly Asp Gly Pro Ile Cys Leu Val Leu Ala Pro Thr
165 170 175
Arg Glu Leu Ala Gln Gln Val Gln Gln Val Ala Ala Glu Tyr Cys Arg
180 185 190
Ala Cys Arg Leu Lys Ser Thr Cys Ile Tyr Gly Gly Ala Pro Lys Gly
195 200 205
Pro Gln Ile Arg Asp Leu Glu Arg Gly Val Glu Ile Cys Ile Ala Thr
210 215 220
Pro Gly Arg Leu Ile Asp Phe Leu Glu Cys Gly Lys Thr Asn Leu Arg
225 230 235 240
Arg Thr Thr Tyr Leu Val Leu Asp Glu Ala Asp Arg Met Leu Asp Met
245 250 255
Gly Phe Glu Pro Gln Ile Arg Lys Ile Val Asp Gln Ile Arg Pro Asp
260 265 270
Arg Gln Thr Leu Met Trp Ser Ala Thr Trp Pro Lys Glu Val Arg Gln
275 280 285
Leu Ala Glu Asp Phe Leu Lys Asp Tyr Ile His Ile Asn Ile Gly Ala
290 295 300
Leu Glu Leu Ser Ala Asn His Asn Ile Leu Gln Ile Val Asp Val Cys
305 310 315 320
His Asp Val Glu Lys Asp Glu Lys Leu Ile Arg Leu Met Glu Glu Ile
325 330 335
Met Ser Glu Lys Glu Asn Lys Thr Ile Val Phe Val Glu Thr Lys Arg
340 345 350
Arg Cys Asp Glu Leu Thr Arg Lys Met Arg Arg Asp Gly Trp Pro Ala
355 360 365
Met Gly Ile His Gly Asp Lys Ser Gln Gln Glu Arg Asp Trp Val Leu
370 375 380
Asn Glu Phe Lys His Gly Lys Ala Pro Ile Leu Ile Ala Thr Asp Val
385 390 395 400
Ala Ser Arg Gly Leu Asp Val Glu Asp Val Lys Phe Val Ile Asn Tyr
405 410 415
Asp Tyr Pro Asn Ser Ser Glu Asp Tyr Ile His Arg Ile Gly Arg Thr
420 425 430
Ala Arg Ser Thr Lys Thr Gly Thr Ala Tyr Thr Phe Phe Thr Pro Asn
435 440 445
Asn Ile Lys Gln Val Ser Asp Leu Ile Ser Val Leu Arg Glu Ala Asn
450 455 460
Gln Ala Ile Asn Pro Lys Leu Leu Gln Leu Val Glu Asp Arg Gly Ser
465 470 475 480
Gly Arg Ser Arg Gly Arg Gly Gly Met Lys Asp Asp Arg Arg Asp Arg
485 490 495
Tyr Ser Ala Gly
500

<210> 14
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> protein domain

<400> 14

Thr Gly Ser Gly Lys Thr
1 5

<210> 15
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> protein domain

<400> 15
Asp Glu Cys His
1

<210> 16
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> protein domain

<400> 16
Ala Arg Gly Arg Ala
1 5

<210> 17
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> protein domain

<221> VARIANT
<222> 8
<223> Xaa = Any Amino Acid

<221> VARIANT
<222> 9
<223> Xaa = Any Amino Acid

<400> 17
Tyr Ile His Arg Ile Gly Arg Xaa Xaa Arg
1 5 10

<210> 18
<211> 8
<212> PRT
<213> Artificial Sequence

<220>

<223> protein domain

<400> 18

Leu Pro Thr Gly Ser Gly Lys Thr
1 5

<210> 19

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> protein domain

<221> VARIANT

<222> 1

<223> Xaa = Ala or Gly

<221> VARIANT

<222> 2

<223> Xaa = any amino acid

<221> VARIANT

<222> 3

<223> Xaa = any amino acid

<221> VARIANT

<222> 5

<223> Xaa = any amino acid

<400> 19

Xaa Xaa Xaa Gly Xaa Gly Lys Thr
1 5

<210> 20

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> protein domain

<400> 20

Asp Glu Ala Asp

1

<210> 21

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> protein domain

<400> 21
Asp Glu Ala His
1

<210> 22
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> protein domain

<221> VARIANT
<222> 3
<223> Xaa = Any Amino Acid

<400> 22
Asp Glu Xaa His
1

<210> 23
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> protein domain

<221> VARIANT
<222> 6
<223> Xaa = Any Amino Acid

<221> VARIANT
<222> 7
<223> Xaa = Any Amino Acid

<400> 23
His Arg Ile Gly Arg Xaa Xaa Arg
1 5

<210> 24
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> protein domain

<400> 24
Ala Arg Gly Arg Ile
1 5

<210> 25

<211> 6406

<212> DNA

<213> homo sapiens

<400> 25

tccactcaat ataaagcttg cactcattct ccaagccccag gtgtgatccg attcttccag 60
tataccaaagt caagaacacctg ggatacagaa agccctctgt ccttgagaca atgttagaggg 120
tctaactgag ctgttaaca caagccacct atagacagca aaactaaaag atcaccctgt 180
aacacacgccc cactgaggct tcagaagctg taaacatcca cccctagaca ctgcccgtggg 240
tcggagcccc acagcctgccc catctgcagg ctcccctaga ggttgagca gtggggact 300
gaagaagcga gccacaccccc catactgccc aaggtaattt acagattcaa tgccatcccc 360
atcaagctac caatgacttt ctgcacagaa ttggaaaaaa ctactttaaa gttcatatgg 420
aaccaaaaaaaaa gagcccgcat cgccaaagtca atcctaagcc aaaagaacaa agctggaggc 480
atcaccctac ctgacttcaa acaataactac aaggctacag taacccaaac agcatggtac 540
tggtacccaa acagagatat agatcaatgg aacagaacag agccctcaga aataatgcca 600
catatctaca actatctgt ctgtacaaa cctgagaaaa acaagcaatg gggaaagtat 660
tccctatTTtta ataaatggtg ctggggaaaac tggctagcca tatgtagaaa gctgaaactg 720
ggttccttcc ttacacctta tacaaaaatc aattcaagat ggattaaaga cttaaacgtt 780
agacctaaaaa ccataaaaaac cctagaagaa aacctaggca ttaccattca ggacatacgc 840
atgggcaagg acttcatgtc taaaacacca aaagcaatgg caacaaaagc caaaattgac 900
aaacgggtatc taattaaact aaagagcttc tgcacagcaa aagaaaactac cattagagt 960
aacagggcaac ctacaaaatg ggagaaaattt ttcgcaacct actcatccga caaagggcta 1020
atatccagaa tctacaatga actcaacaa atttacaaga aaaaaacaaa caaccccatc 1080
aaaaagtggg tgaaggacat gaacagacac ttgtcaaaag aagacattta tgcagccaaa 1140
aaacacatga aaaaatgctc accatcaactg gccatcagag aaatgcaat caaaaccaca 1200
atgagatacc atctcacacc agttagaatg gcaatcatta aaaagtcaagg aaacaacagg 1260
tgcgtggagag gatgtggaga aatagaaca cttttgcact gttggggaa ctgtaaacta 1320
gttcaaccat tgcgtggatc agtgcgtga ttctcaggg atctagaact agaaatacca 1380
tttgaccctg ccatccatt actgggtata tactcaaagg actataaattc ttgtctgtat 1440
aaagacacat gcacatgtat gtttattgtg gcattattca caatagcaaa gacttggAAC 1500
caacccaaat gtccaaacagt gatagactgg attaagaaaa tgcgtggacac atacaccatg 1560
gaataactatg cagccataaa aatgatgag ttcatgtcct ttgttagggac atggatgaaa 1620
ttggaaatca tcattctcag taaactatcg caagaacaaa aaaccaaaca cccgcatattc 1680
tcactcatag gtggaaattt aacaatgcga acacatggac acaggaagga gaacatcaca 1740
ctctggggac ttttgggggg tggggggagg gggggggat agcattggta gatataccta 1800
atgcgtatg acgagttatg gggtgcagcg caccagcatg acacatgtat acatatgtaa 1860
ccaacctgca cattgtgcac atgtacccta aaacttaaaag tataataata aataaataaa 1920
taaataaataa aataaataaa gtaaaataaa acaattacaa tctagcctt gaggtaaaag 1980
tactgtttt cacaaaaaaca tttgcaggtt actgtttttt aaaagacttt aagctatgga 2040
aggagtactt gaaaaatgaa ttttccaaaaa cttatctatt gatacgtgac tttcattttt 2100
tgccaaaact gctatgtaga aaagttttt tatgtgaaaa cttaaaaacc agaattttaa 2160
ttgaatttggt gaaagtgtatt agggaaattat tatcaagatt tagtgaactt agccataatt 2220
ttttttctat tttaggctta ctactatTTtta tggaaataaaa agctacgaca gtatccttt 2280
aataaacttt cctgctaaat cagccatca gtttcagttt aatggctgaa agtcttgctt 2340
aaagtctcag taaaatggct agctattata tagtgcattt atgtatgtgt gtatataat 2400
atatatataat atatatataat atatatataat atatatataat atgttaactaa atttttcctt 2460
tataaattgt gcattcttg aagactagca ccgcaccatc ttttctttaa tttttatata 2520
agcgttagtgg gctggagtca catattgggc acataaacat gccaggctgg tgcttagtgc 2580
ttacagtcta tccttagaaac aaacttctga catgatacca gaatcttcc attttacaac 2640
tgatgttatt gaggtgattt ttcaaaagcac agcaattaag aaatagtatt gagatgtgaa 2700
ctcagacagc ctgaactcag agtctctgtt cttaaccata ccccacactg ccaggttaag 2760
agcatctaac actttaaattt acacaaagca ggctcattat tgatacaat gacaaacaaa 2820
gtaaaggAAC agaacaacaa ttccagggtt tctcaactaa cttaaaattat tgcattttc 2880
tttggaaaaag acattattgc tatgcattgtt cgtttaattt tagtggcagc tcatattgtt 2940
actacttctt aaaaactcaa atgaaaaggat gctataacaat gggaaaatac atagttcagc 3000
aggatctcct gcctcaaaag agaaaggaaa aagaaactta cattggggaa ctgggtaaaa 3060
qqattaaaaat qaaacctaqt agaaqaaaact tqacacaqqqa aaacaattaa ttactcaagt 3120

gaaaaacaga aaataaaacta aatcatgatg caaaaaatat agatgaaaaa aggatacatt 3180
gtgagagatt gtgtcttgc ttttgccttcc ttaacccctt ttctccaaaa agggcccat 3240
caagactatg ggagattcct aaaaaagaag tcccttcac ccacacctaa tcctcatcac 3300
tcagaccta tccagcagag agactcctac ttgtgagaaa atatgaattt gtttgcgttgg 3360
gtattatgtg atgctaatac ggttagagga ggtactat ttggaaatc aacctgtgaa 3420
actgtatatac acattattat gtagatttac tatggcttc agggcattt tcctcacctg 3480
cacattgcat attttttagt cattactac catctatctt cccactccca tttagaatgtg 3540
aactccataa aagttaggagc ttgttaatt ttattactg caccatgatc agtgctggc 3600
atgtataga tacttaataa acatattttt aatgactaga tgatacaatg aatgatataa 3660
tttgaatgcc aaatatttaa atatcttgg tttaaatgtt tattatttga gaacaggc 3720
aataccaaac atttgatcct ttcttccca gagcaacaat taagtggat gaagaaaata 3780
acattaactg gttccccata ttcagtcgc aacccttctt cattccccc tgtttgaac 3840
caagaaacag aaggataagt gccaggaaaa agaatgttt ttgggttggt agtttgcgt 3900
ttaacatgtt gaataaaacc cactggcgc tgggggatag gagatgtt ttgcaacagc 3960
cttaaaaatg attttcatag acccaatact taaaattaat aatttgatgt cttttgtaga 4020
aacatctaat tggattatcc tctatctggg tacaaggc tctccaaaaa ttaagtgaaa 4080
aaggagtagg gttctgtgaa gaaacagaaaa agaacagat aatcaggcc tacctgc 4140
cccaagggtt catttacttc aactttcagt gtatttaaca ttatgcccgc tgctatggc 4200
aactcaaata caactcccgag gagagatgtc atcaagagcc caccagggtt gagtagtgta 4260
ctagttacta tgtaaatata tcccttctt aagcagctt caatccttca ggggttagaaa 4320
aagccttgct acataagata attagagaat aaaataagac atgttaccat aaagtgc 4380
tttggatatt ttgtatgccc cttagaaaca ctcaccaaga ctctgtactt ctattatctt 4440
gttcaagca ctatcaggtt tccctggcct acacagactt tattgaatgt actttgc 4500
cagattattt ttccctaaata tgcctcttgc taaacctaaa tgatctcttcc atccttata 4560
taattctgga ccatgagatt ctatgtatgg tgcgtatgtg cctaccaccc acagtcacat 4620
gtggctacag aatgccttca gaatgagtag taaccttaag gactcacatt tatgtggc 4680
ctgtacccaa atgaagctgc catttttcag tgcgtatgtg tttttttct ctcatgacat 4740
agacaaatgt tgatgtttac tacaagttgg tacatttagt gctaattaag ttccctagctg 4800
ctccagccaa aacttgcgtt attgaatcca agaaaaagaat ggcagctata tcaaaaataa 4860
gttgggggg gatttttttt ttttgcctt taaaggaaa gttgtatatt aaagaatata 4920
gggaacttac aagctgggat cttagaaact taaagtcgg gttcccttct aagctgagtt 4980
ggtggttcaat gtcctccac atctgttacc aggtcctgtt caagctgc taaataccag 5040
caatctaaat atgaggcagt aaagttaaact gtttattgtt actcaactt tcgaacccac 5100
ctccaaattt ccagggaaac aagttgttgc ttggaaaccc acaggaggc aggtttattt 5160
taggaaggac ttccctctgt ctctccaca tctctgc 5220
tctcacctgt ccctcgcatg ctacccaccc tcagccaggc ctgcctacat tcaccagccg 5280
agggttaactc cctgttcaacg tccgggtctg tggcgttgc ttttgccttcc ccccttggaa 5340
agtccaaat cacatgttt tatgcctgc acatgttgc ctacaaagga ctttattgtt 5400
aaggcagaac ctgctggaa aacaaatat ccggccggagg agcttgc tgcgttgg 5460
cttgggttca gagagaattt gcttccctt tctgtttccc ggggtgttca taaaccagg 5520
ctccctcttca tccccggccc cgacccaaag gtggcgttgc cctgaggaaa cccctcccc 5580
gccaggcaga ttacgtttac aaagtcctga gaagagaatc gaaacagaaa ccaaagtgc 5640
gcaaactctg taagaactgc ctgacagaaa gctggactca aagctcttac ccgagtggtc 5700
agcaggatcg ccccggttccg ggacccagg cgcacaccgc agagtc 5760
tgccggccgc acctgcctgc cgcggcccg cgcggccgc cgtgcccac ctgcccgc 5820
gcccacctgc ccaggtgcga gtgcagcccc ggcggccggc ctgagagccc tgggacaac 5880
ctcgatcttgc tcaggcatac agcggttagac cctgcttctc taagtggca gggacagc 5940
gcacgcacat ttacactgtc ccgcagacaa cagcaccatc tgcttggag aaccctctt 6000
cttctctgag aaagaaaatg gtcgaatggg tattccacag acgagaattt cgcctatctc 6060
atctcgatgttca tcaaggccat ggtgaaaatg tacatccagg tggagctgt gctggactac 6120
ctgaccccttca tgcctgcaga ggtgaaggag cagattcaga ggacagtcgc cacccggg 6180
aacatgcagg cagttgaact gtcgttgc accttggaga agggagtcg gcaccccttgc 6240
tggactcggg aattcgtgga gcccctccgg agaaccggca gccccttgc cggccgc 6300
atgaacccttgc agtcacggc ttgccttca ccatgttttgc agaaccgtca tgcgttgc 6360
ctccaaacttgc tgaacccttca tcaaggccact ctgggtggaca agcttc 6406